

## Sources of Tropical Subseasonal Skill in the CFSv2

CARL J. SCHRECK III

*North Carolina Institute for Climate Studies, North Carolina State University, Asheville, North Carolina*

MATTHEW A. JANIGA

*Marine Meteorology Division, Naval Research Laboratory, Monterey, California*

STEPHEN BAXTER

*NOAA/Climate Prediction Center, College Park, Maryland*

(Manuscript received 30 August 2019, in final form 22 January 2020)

### ABSTRACT

This study applies Fourier filtering to a combination of rainfall estimates from TRMM and forecasts from the CFSv2. The combined data are filtered for low-frequency (LF,  $\geq 120$  days) variability, the MJO, and convectively coupled equatorial waves. The filtering provides insight into the sources of skill for the CFSv2. The LF filter, which encapsulates persistent anomalies generally corresponding with SSTs, has the largest contribution to forecast skill beyond week 2. Variability within the equatorial Pacific is dominated by its response to ENSO, such that both the unfiltered and the LF-filtered forecasts are skillful over the Pacific through the entire 45-day CFSv2 forecast. In fact, the LF forecasts in that region are more skillful than the unfiltered forecasts or any combination of the filters. Verifying filtered against unfiltered observations shows that subseasonal variability has very little opportunity to contribute to skill over the equatorial Pacific. Any subseasonal variability produced by the model is actually detracting from the skill there. The MJO primarily contributes to CFSv2 skill over the Indian Ocean, particularly during March–May and MJO phases 2–5. However, the model misses opportunities for the MJO to contribute to skill in other regions. Convectively coupled equatorial Rossby waves contribute to skill over the Indian Ocean during December–February and the Atlantic Ocean during September–November. Convectively coupled Kelvin waves show limited potential skill for predicting weekly averaged rainfall anomalies since they explain a relatively small percent of the observed variability.

### 1. Introduction

Subseasonal-to-seasonal (S2S) forecasting remains one of the most challenging but potentially beneficial horizons in atmospheric science. These forecasts are largely driven by the divergent circulations associated with tropical convection, which can broadly be categorized in terms of their temporal periods and/or zonal wavelengths. The MJO in particular has a period of 30–60 days and is an ideal target for these types of forecasts (Zhang 2013). Higher-frequency modes like convectively coupled Kelvin waves may have shorter-term effects at any given location, but they circumnavigate the globe over the course of a month on S2S time scales (Straub et al. 2006; Ventrice et al. 2012b,a; Ventrice

and Thorncroft 2013; Schreck 2015, 2016). Similarly, while ENSO has a longer period than the MJO, its transitions can fall squarely in the S2S window and it modulates the activity of higher-frequency modes (Kessler 2001). Model skill for predicting S2S tropical variability has improved over recent years, but it still fails to leverage the full predictive potential of these tropical forcings (Gottschalck et al. 2010, 2013; Hamill and Kiladis 2014; Oliver and Thompson 2015).

Wheeler and Weickmann (2001) developed a method for filtering proxies for convection like OLR in near-real time to identify the MJO and convectively coupled equatorial waves (CCEWs). Fourier filtering breaks down at end points, which poses a challenge for real-time applications. Wheeler and Weickmann overcame this obstacle by padding the observed OLR data with climatology. The resulting filtered anomalies extrapolated

---

*Corresponding author:* Carl J. Schreck III, carl\_schreck@ncsu.edu

DOI: 10.1175/MWR-D-19-0289.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

the current state of the MJO and CCEWs with skillful forecasts out to 15–20 days for the MJO.

Schreck (2013) extended the Wheeler–Weickmann (Wheeler and Weickmann 2001) method to support forecast operations during the Dynamics of the MJO (DYNAMO) field campaign (Yoneyama et al. 2013; Gottschalck et al. 2013) by padding the observations with 45-day forecasts from CFSv2. This combination should be more skillful than the Wheeler–Weickmann method because the model forecasts are generally more skillful than climatology.

Janiga et al. (2018) showed that this method provides a valuable way to evaluate the ability of the model to simulate tropical variability and determine the sources of predictive skill. Compared with other metrics like the Wheeler–Hendon Real-time Multivariate MJO (RMM; Wheeler and Hendon 2004) index, Fourier filtering is better at identifying regional variations in skill associated with the MJO and CCEWs. Janiga et al. (2018) evaluated control member forecasts of OLR and zonal winds at 850 and 200 hPa from the Navy Earth System Prediction Capability (ESPC), the ECMWF, and the CFSv2 models during boreal summer 1999–2015. Overall, the ECMWF model had the smallest biases in MJO and CCEW activity. Navy ESPC had a bias toward overamplifying the MJO and CCEWs. The CFSv2 was generally the opposite; it underestimated the amplitude of the MJO and convectively coupled Kelvin waves.

One of the biggest contributions of Janiga et al. (2018) was identifying the primary sources of skill within the model. For example, the low-frequency (LF; >100 days) filter had a lower mean absolute error (MAE) than the unfiltered forecasts for all three models beyond 1. Adding the MJO-filtered anomalies to the LF-filtered anomalies substantially improved the anomaly correlation coefficient (ACC) for week 1 with decreasing benefits for subsequent weeks. Meanwhile, equatorial Rossby (ER) waves only improved the ACC in week 1, and other CCEWs made no significant contribution to the skill. Taken together, these results suggest that models derive much of their skill from simply maintaining the LF state with the MJO contributing to some skill in weeks 1–2. Meanwhile, the higher-frequency variability actually detracts from model skill after one week.

This study builds upon Janiga et al. (2018) by examining CFSv2 skill during all seasons using the ensemble mean. We also examine variations in skill between regions and by MJO phase. While CFSv2 generally had the lowest skill in their study, it remains the model that is most accessible to forecasters at NOAA/CPC and elsewhere. A thorough understanding of its strengths and weaknesses can therefore be useful to those forecasters.

To that end, we also focus on Heidke skill score (HSS) for verification, since that is the preferred metric at CPC. The reference used to evaluate the skill of the filtered forecasts is unfiltered anomalies since this also provides information on the contribution of different filters to the total variability.

## 2. Data and methods

### a. Data

This study examines CFSv2 skill for rainfall 1999–2010. The model forecasts are initialized every 6 h and output at 6-hourly timesteps out to at least 45 days (Saha et al. 2010). For this study, the 6-hourly timesteps were averaged to create daily forecasts, and the four initializations each day were averaged to create a four-member ensemble mean. About 2% of the daily ensembles were omitted because of missing or corrupted data. Model climatology and biases were removed by subtracting the 1999–2010 CFSv2 reforecast calibration climatologies.

The CFSv2 was verified against TRMM multisatellite precipitation analyses (TMPA, TRMM3b42; Huffman et al. 2007). The verification was performed on the tropical strip 20°S–20°N. The TRMM data were averaged from 3-hourly to daily, and both the TRMM and CFSv2 datasets were regridded to the same 1.25° latitude–longitude grid. Anomalies were calculated by subtracting the first three harmonics of the annual cycle for 1998–2015.

### b. Filtering methods

The Fourier filtering methodology closely follows that from Janiga et al. (2018). Figure 1 shows an example of the filtering method. The reference filtering (Fig. 1a) applies the filters to the entire TRMM dataset (1998–2015) as is typically done in observational studies (Kiladis et al. 2005, 2009). The observational padding (Fig. 1b) uses 411 days of TRMM observations followed by 655 days of zero anomalies (climatology). This arrangement is similar to the method from Wheeler and Weickmann (2001). They found that the filter has very little sensitivity to the lengths of each segment of the data. One difference is that here the observed anomalies are used for an additional 45 days to emulate a perfect forecast. This method illustrates the contribution of each filter to the observed variability and the potential utility of each mode for forecasting rainfall. It also shows the sensitivity of the filters to the padding. Differences in the filtered contours between the reference filtering (Fig. 1a) and the observational padding (Fig. 1b) become evident after 13 December (28 days) in this example as the filter increasing responds to the climatological padding.

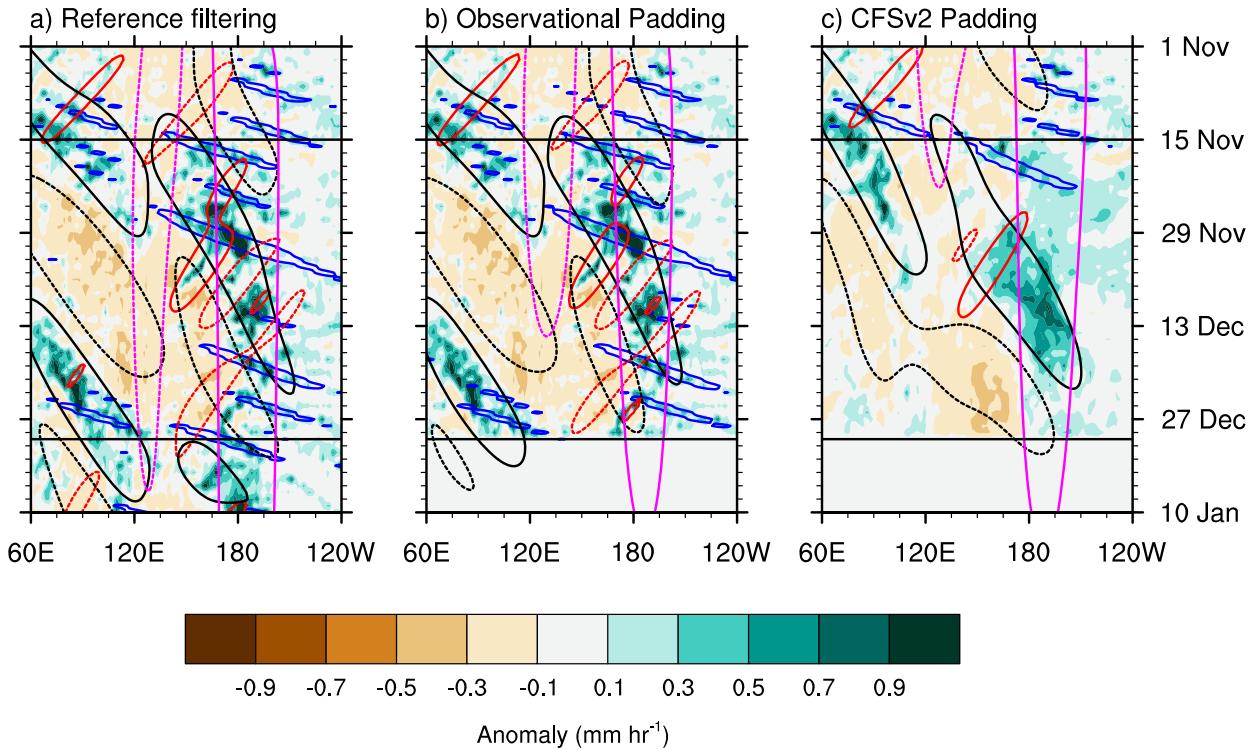


FIG. 1. Hovmöllers of a sample forecast initialized on 15 Nov 2002. Shading indicates unfiltered rainfall anomalies and contours identify LF (magenta), MJO (black), ER-wave (red), and Kelvin wave (blue) filtered anomalies. These filters are applied to three sets of data: (a) the entire TRMM observational record (1998–2015); (b) observed TRMM anomalies until 30 Dec 2002 (lower horizontal line) and climatology (zero anomalies) thereafter; and (c) observed TRMM anomalies until 15 Nov 2002 (upper horizontal line), CFSv2 forecasts for 15 Nov to 30 Dec 2002, and climatology (zero anomalies) thereafter. The reference filtering in (a) is shown only for reference. The observational padding in (b) is used in Figs. 2–4, and the CFSv2 padding in (c) is used for all other figures.

The primary focus of this study is on the CFSv2 filtering (Fig. 1c). In this case, the filtering is applied to an array that contains 366 days of TRMM observations followed by 45 days of CFSv2 forecasts and 655 days of zero anomalies (climatology). The differences between the CFSv2 forecasts and the observed rainfall anomalies affects the filtered anomalies even before the initialization date (see differences between the contours in Figs. 1b and 1c during 1–15 November). However, these differences are small and decrease to zero around a week before forecast initialization [Fig. 3 of Janiga et al. (2018)] but this study will only examine the filtered anomalies during the forecast period (15 November to 30 December in this example). Both the observational and CFSv2 paddings are repeated for each day during 1999–2010.

Based on the results of Janiga et al. (2018), this study focuses primarily on three filter bands: low-frequency/large-scale (LF, wavenumber  $k = -10:10$ ,  $p > 120$  days), the MJO [ $k = 0:9$ ,  $p = 20\text{--}100$  days, following Kiladis et al. (2005)], and ER waves [ $k = -10:-1$ ,  $p = 10\text{--}72$  days,  $h < 90$  m, following Kiladis et al. (2009)] where  $k$  is the

zonal wavenumber,  $p$  is period, and  $h$  is the shallow water equivalent depth. We will also briefly compare two Kelvin wave filters: one from Kiladis et al. (2009) that is based on the shallow water dispersion lines ( $k = 1:14$ ,  $p = 2.5:30$  days,  $h = 8:90$  m) and a broader rectangular filter ( $k = 1:14$ ,  $p = 2.5:20$  days) (Janiga et al. 2018; Roundy 2012a).

Following Janiga et al. (2018), the LF-filtered anomalies are subtracted from the nonzero portion of the data before filtering for the MJO and ER waves. Subsequently, a 20-day low-pass filter is subtracted from the nonzero data before filtering for the Kelvin waves. These steps reduce some of the spectral ringing that can occur at the interface between the zeros and nonzeros.

c. Heidke skill scores (HSS)

One of the primary components of CPC’s Global Tropical Hazards and Benefits Outlook is forecasts for above and below average precipitation as defined by tercile categories. The Heidke skill score (HSS) is an ideal metric for verifying such forecasts (O’Lenic et al. 2008; Sooraj et al. 2012; Peng et al. 2013). HSS provides

the percentage of forecasts that match the observed category relative to what might occur just by chance. It is defined as

$$\text{HSS} = 100 \times \frac{H - E}{T - E},$$

where  $H$  is the number of correct forecasts (hits),  $E$  is the expected number of correct forecasts by chance, and  $T$  is the total number of forecast–observation pairs. For tercile forecasts,  $E = T/3$ . A perfect forecast would score 100 while a forecast worse than random chance would score  $<0$ . All HSS calculations use daily initializations of the forecasts from 1999 to 2010.

In this study, terciles are calculated separately for the TRMM observations, the unfiltered CFSv2 forecasts, and for each filter band. They are calculated at each grid point and for each calendar month using data from all years (1999–2010). The forecast terciles are also calculated for each lead. Both filtered and unfiltered forecasts are thus compared to their own terciles to determine the forecast tercile, which is then verified against the observed unfiltered TRMM tercile.

Statistical significance of the HSS results is evaluated using the cumulative density function for a binomial distribution (Wilks 2006, p. 74). The effective sample size was evaluated following [Wilks 2006, 144–145, Eq. (5.12)]:

$$n' \cong n \frac{1 - \rho_1}{1 + \rho_1},$$

where  $n'$  is the effective sample size,  $n$  is the number of forecasts being verified, and  $\rho_1$  is the lag-1 autocorrelation coefficient of the unfiltered TRMM rainfall observations at a given point. The cumulative density function is then used to estimate how many hits would occur from random chance (33% for tercile forecasts) at the 95% level. The  $\text{HSS}_{95}$  is then calculated based on this number of hits and effective sample size. The actual HSS is considered statistically significant if it is greater than  $\text{HSS}_{95}$ .

A similar method is used to evaluate the statistical significance of the difference in HSS between two sets of forecasts. The cumulative density function is used to calculate the 95% range of HSS for each set of forecasts given their hit rates ( $H/n$ ) and the effective sample  $n'$ . The difference is considered statistically significant if either actual HSS falls outside of the 95% range of HSS for the other.

### 3. Results

Figure 2 shows the HSS between observational padding filtered TRMM rainfall anomalies and the unfiltered anomalies at week 3. Verifying observed filtered

anomalies against unfiltered ones illustrates the contribution of each filter to the observed variability and the potential utility of each mode for forecasting rainfall. The results are only nominally sensitive to lead through week 3, but thereafter the filters begin to lose skill as they are affected more by the beginning of the zero padding (cf. Figs. 1a,b).

The LF filter (Fig. 2a) explains the largest amount of variability over the western Pacific where ENSO-related SST anomalies strongly control the rainfall. Meanwhile, the MJO's impact (Fig. 2b) is largest over the Indian Ocean and to a lesser degree the Maritime Continent. The variability associated with ER waves (Fig. 2c) is comparable to that of the MJO over most of the tropics. The primary exceptions are that ER-wave variability is somewhat weaker than the MJO over the Indian Ocean, and it is virtually absent over the Maritime Continent.

The MJO and ER waves explain little of the rainfall variability over the equatorial Pacific, even though they have large variance there (Kiladis et al. 2005, 2009). This difference illustrates one of the key limitations in using subseasonal modes alone to predict rainfall. Bandpass-filtered anomalies are required to sum to zero over one full wavelength and/or period. Positive anomalies must be coupled with subsequent negative anomalies and vice versa. This becomes problematic when there are strong LF anomalies. For example, the forecast in Fig. 1 is taken from the strong El Niño during 2002/03. That El Niño favored positive rainfall anomalies near the date line (solid magenta contours). The dry phase of the MJO in late December (dashed black contours) reduces those positive anomalies (green shading) but produces hardly any dry anomalies (brown shading). As a result, verifying the MJO-filtered anomalies alone against the unfiltered anomalies would find little skill.

Figures 2d and 2e show the HSS for the sum of the LF and MJO (LF+MJO), as well as the ER waves (LF+MJO+ER). The skill increases with the addition of each mode. Those increases are further quantified in Fig. 3 by comparing the HSS with and without each sum. The improvements to skill from adding the MJO to the LF (Fig. 3a) largely align with the regions where the MJO filter produces skill on its own (Fig. 2b). Adding the MJO detracts from skill over the equatorial and Southeast Pacific, but those decreases are not statistically significant (gray shading, see section 2c). Adding ER-filtered anomalies improves skill (Fig. 3b), but those improvements over LF+MJO alone are fairly modest (Fig. 3c). Most of those improvements occur in regions where tropical cyclone activity commonly projects onto the ER filter band (Schreck et al. 2012; Ayyer et al. 2012).

Kelvin waves are another potential source of subseasonal skill. Their lifespans often extend into the

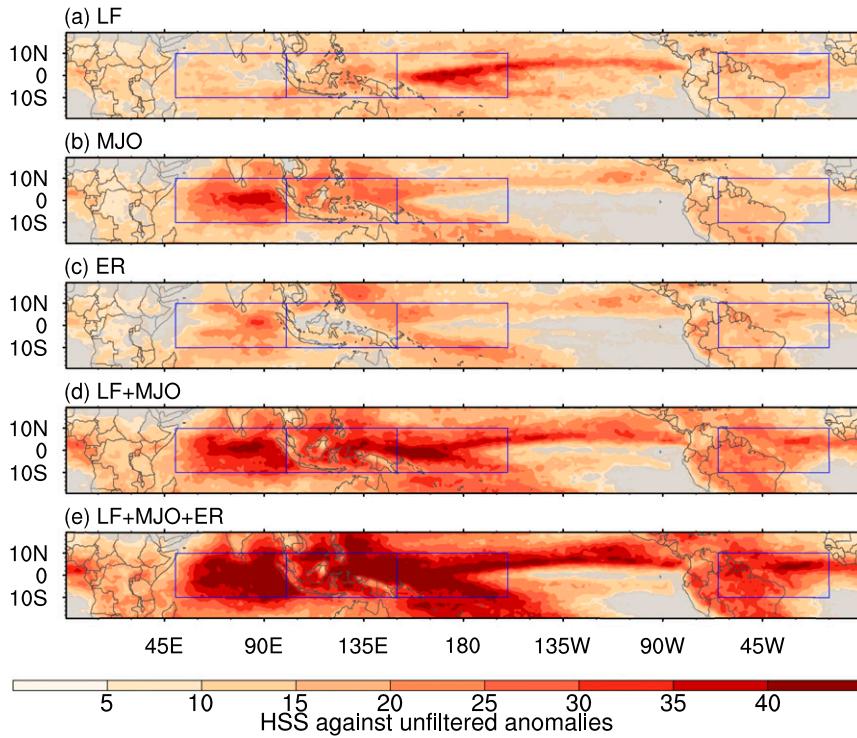


FIG. 2. HSS between observed filtered rainfall anomalies and observed unfiltered rainfall anomalies for times corresponding to week 3 forecasts for (a) LF, (b) MJO, (c) ER-wave filters, (d) sum of LF and MJO filtered anomalies, and (e) sum of LF, MJO, and ER-filtered anomalies. Gray shading denotes areas where HSS is not significantly different from zero (see section 2c for details). Boxes outline primary regions of skill that will be examined in Figs. 7–9. From west to east, they are the Indian Ocean, Maritime Continent, western Pacific, and Atlantic.

subseasonal scale (Straub et al. 2006), and they may exist on a continuum with the MJO (Roundy 2012b,c; Sobel and Kim 2012). They are often identified with a fairly narrow filter band bounded by the shallow water dispersion curves

for equivalent depths of 8 and 90 m (Wheeler and Kiladis 1999; Kiladis et al. 2009). Although this filter captures largest spectral peak, it misses many other eastward propagating systems that share Kelvin characteristics, so

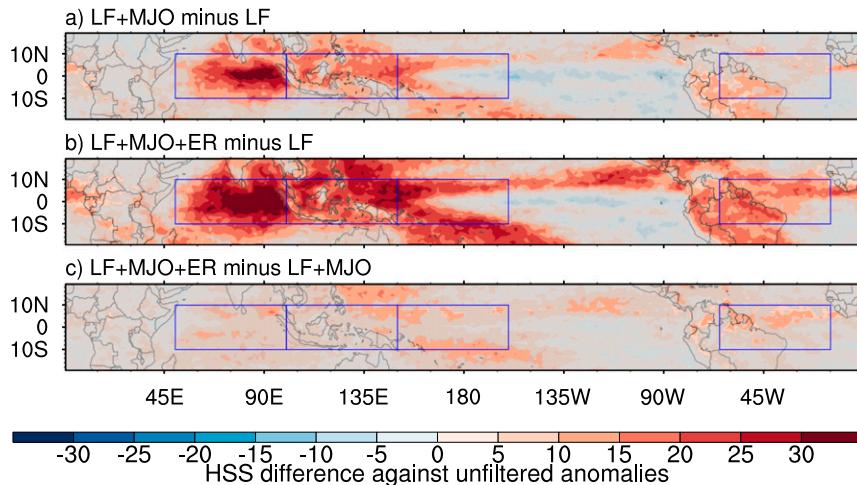


FIG. 3. Differences in HSS for various combinations of filtered rainfall anomalies verified against unfiltered rainfall anomalies. Gray shading denotes areas where that difference is not significantly different from zero (see section 2c for details).

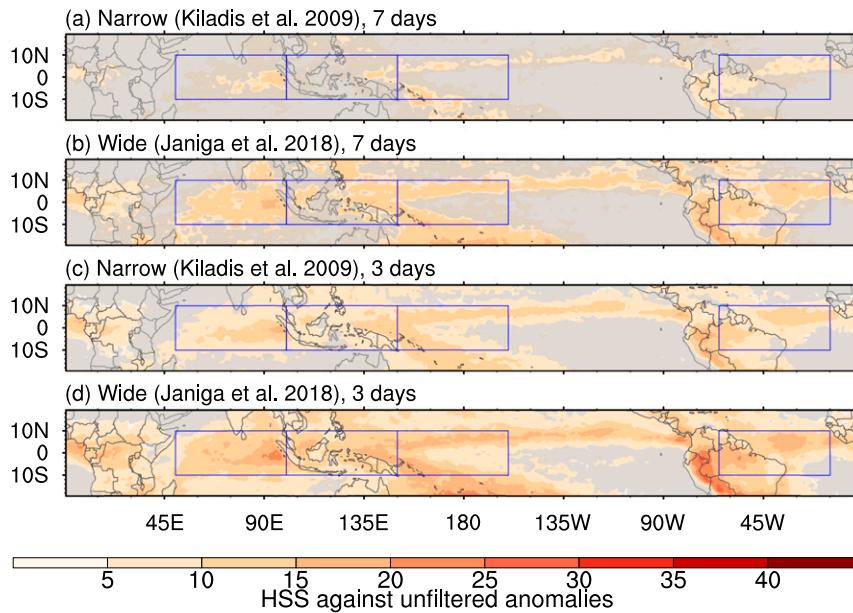


FIG. 4. As in Fig. 2, but for Kelvin waves using either (a),(c) a narrow filter from Kiladis et al. (2009) or (b),(d) a broader rectangular filter from Janiga et al. (2018) and either averaging over (a),(b) 7 days or (c),(d) 3 days.

broader rectangular filters are also often used (Roundy 2012a; Janiga et al. 2018). The size of the filter comes with trade-offs. A narrow filter can extrapolate the existing signals farther (Wheeler and Weickmann 2001), but it naturally reduces the variance within the filter. A broader filter, on the other hand, can more faithfully approximate the unfiltered anomalies, but it also depends more strongly on the skill of the underlying model forecast data.

Figure 4 illustrates the amount of the total rainfall variability explained by these types of filters. The narrower Kiladis et al. (2009) filter only explains a small portion of the variability when averaged over weekly periods (Fig. 4a), which makes sense because the filter is particularly narrow at longer periods. The wider filter explains more of the variability of weekly averaged rainfall (Fig. 4b), although still less than the other filter bands (Fig. 2). For 3-day windows the narrow Kelvin filter explains more of the rainfall variability (Fig. 4c), but still not as much as the wider Kelvin filter (Fig. 4d). We will not consider Kelvin waves further in this study because of their small potential skill and the fact that the CFSv2 underrepresents these waves (Goswami et al. 2017; Janiga et al. 2018). However, future studies should continue to investigate ways to harness subseasonal skill from these waves.

Figure 5 shows the week-3 HSS scores for the CFSv2. The unfiltered CFSv2 rainfall anomalies (Fig. 5a) show skillful forecasts across most of the tropical ITCZ.

The strongest skill is in the Pacific, particularly near the date line. The skill for the LF-filtered CFSv2 anomalies (Fig. 5b) is very similar, aside from reduced skill over the Indian Ocean. This similarity suggests that the CFSv2 derives most of its week-3 skill from essentially persistent features. The patterns are also very similar between the LF-filtered CFSv2 and the LF-filtered observations (Fig. 2a), albeit with the CFSv2 naturally having somewhat lower and more localized skill.

The LF CFSv2 skill (Fig. 5b) is actually larger than that of the unfiltered model output (Fig. 5a) over the equatorial Pacific. This difference implies that the model's higher-frequency (<120 days) variability is detracting from the skill in the equatorial Pacific. In addition, this is only week 3, so the Pacific SSTs will still be close to their initial conditions. The LF skill is not implying that the model is necessarily simulating the evolution of ENSO. Instead, it shows that the model is accurately representing the response to the preexisting ENSO conditions.

The MJO-filtered CFSv2 is only significantly skillful (see section 2c for details of the significance testing) over the Indian Ocean and a small region off the northern coast of Papua New Guinea (Fig. 5c). These are the regions of greatest potential skill (Fig. 2b), but many regions of opportunity are missed, notably the South China Sea, the eastern North Pacific, and Brazil. Meanwhile, the ER-filtered forecasts (Fig. 5d) are not significantly skillful anywhere.

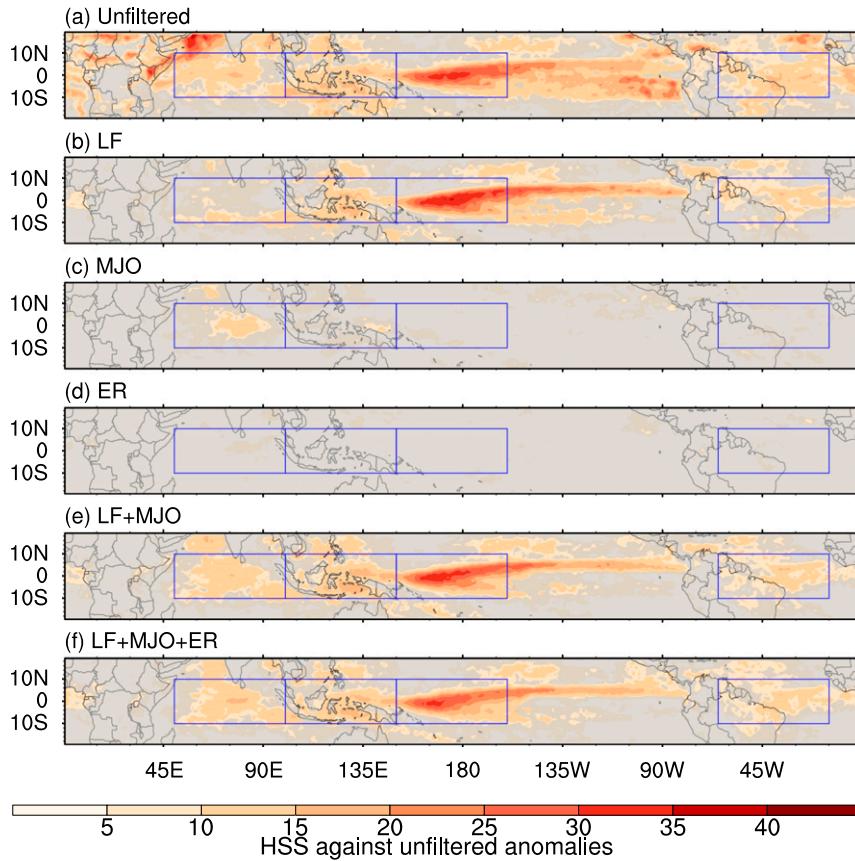


FIG. 5. As in Fig. 2, but for (a) unfiltered and (b)–(f) filtered CFSv2 forecasts.

Despite the lack of skill from the MJO and ER filters alone (Figs. 5c,d), they still contribute to CFSv2’s overall skill. Adding the MJO-filtered anomalies to the LF anomalies (LF+MJO, Fig. 5e) increases the skill over the Indian Ocean above using the LF alone (Fig. 5b). Over the equatorial Pacific, however, the skill is actually decreased somewhat, consistent with observations (Fig. 3a). Adding ER waves (LF+MJO+ER, Fig. 5f) slightly increases the skill over the Indian Ocean and further degrades it over the Pacific.

Figure 6 shows the number of days until the HSS for a 7-day running mean forecast is no longer significantly different from zero. Over the central Pacific, the entire unfiltered 45-day CFSv2 forecast (Fig. 6a) is skillful, as 42 days in Fig. 6 refers to the 7-day running mean for days 39–45. Over most of the tropics, the model is generally skillful out to 14–21 days, with many subregions showing skill to 21–28 days. Africa is a notable exception where the CFSv2 is barely skillful beyond the first week near the Sahel, Zaire, and Angola.

As in Fig. 5, much of the CFSv2’s skill is derived from the LF signal (Fig. 6b), particularly in the equatorial Pacific and the seas within the Maritime Continent.

The LF filter alone gives skillful rainfall forecasts out beyond five weeks in these regions. LF skill over the Indian Ocean is more limited, barely into 2–3 weeks. Adding the MJO-filtered anomalies (Fig. 6c) extends the Indian Ocean skill into week 4, comparable and even slightly better than the unfiltered CFSv2 (Fig. 6a). The MJO-filtered anomalies also extend the skill over Brazil into week 3, although the unfiltered forecasts still retain better skill there. Adding ER-filtered anomalies has a minimal impact on the skill horizon (Fig. 6d).

Figure 7 shows the skill by lead for several regions. Over the entire tropical strip (dashed lines in all panels), the HSS asymptotes to around 6 for the unfiltered forecasts and around 4 for each set of filtered anomalies. The LF filter (red lines) provides the largest proportion of the skill. Including the MJO (blue lines) primarily increases the skill in the first three weeks. ER waves provide some added skill, particularly during the first two weeks. The unfiltered anomalies (black lines) remain the most skillful. The difference between the filtered and unfiltered skill decreases until around week 3, which is also the limit of the added skill from adding the MJO filter (blue lines). It suggests that the MJO and LF

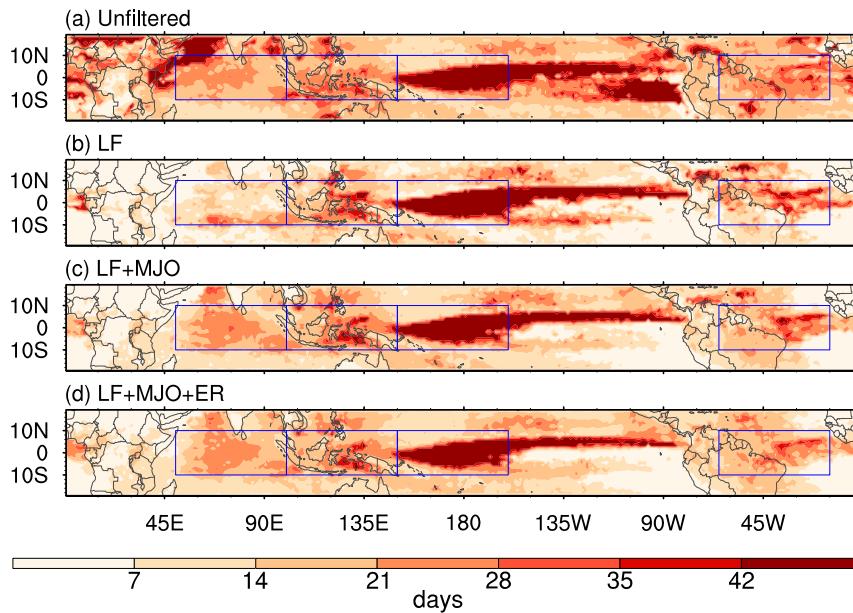


FIG. 6. Number of forecast days until the HSS for 7-day forecast mean is no longer statistically different from zero at the 95% level for (a) unfiltered CFSv2 forecasts and forecasts filtered for (b) LF, (c) LF+MJO, and (d) LF+MJO+ER filtered anomalies. Days shown are the centers of the 7-day window (e.g., day 7 corresponds to the forecast for days 4–10).

variability are the primary driving forces of skill during these three weeks. The unfiltered anomalies probably provide smaller-scale structures associated with these modes that get lost in the filtering. That would explain the higher initial skill of the unfiltered anomalies, but also why that skill diminishes on the same scale as that of the MJO.

The relative contributions of each mode changes between the regions shown in Fig. 7. Not surprisingly, the MJO's impact is greatest over the Indian Ocean (Fig. 7a). Adding the MJO to the LF anomalies nearly doubles the skill in the first two weeks. The unfiltered skill actually falls below that of the LF+MJO in week 3 and beyond. Over the Maritime Continent (Fig. 7b), the skill for LF is about 50% higher over than the entire tropics. However, the MJO also provides added skill through week 3. The skill over the western Pacific (Fig. 7c) is driven largely by the LF signal associated with ENSO. The skill for LF alone in the western Pacific is larger than for any other filter in any of the regions. All of the other filters, including the unfiltered forecasts, degrade the skill after week 2. The Atlantic (Fig. 7d), on the other hand, more closely resembles the tropics at large (dashed lines), albeit with slightly higher skill for all filters.

Figure 8 shows the differences in skill between the LF+MJO and LF alone subdivided by lead time, season, region, and RMM phase and amplitude. Large dots

denote statistically significant differences (see section 2c for details). The season and RMM are for the day of initialization, and the RMM phase is only considered for days when the RMM amplitude is  $>1$ . For the entire tropical strip (top row), adding the MJO generally provides some skill over and above the LF alone (red dots). The improvement is greatest in week 1 and also during March–May. The MJO's impact does not vary much with RMM phase, although it is a little higher when the MJO is in the Indian Ocean to Maritime Continent (phases 3–5). The MJO actually detracts from skill during June–August in phases 2–5. This is consistent with the relatively poor performance of the CFSv2 in simulating the MJO during boreal summer compared to other S2S models (Jie et al. 2017).

The MJO contributions to skill are also divided by RMM amplitude without regard to phase. The amplitudes were roughly divided into terciles by considering days when they were  $<0.9$  (weak),  $0.9$ – $1.5$  (moderate), and  $\geq 1.5$  (strong). As expected, the MJO provides more added skill when the RMM is strong and less when it is weak.

Consistent with Figs. 5e and 7a, the MJO provides its largest improvement to skill over the Indian Ocean (Fig. 8, second row). The improvements are largest during March–May and September–November, where the MJO can improve skill out to week 4. Meanwhile the MJO nominally diminishes the skill over the Indian

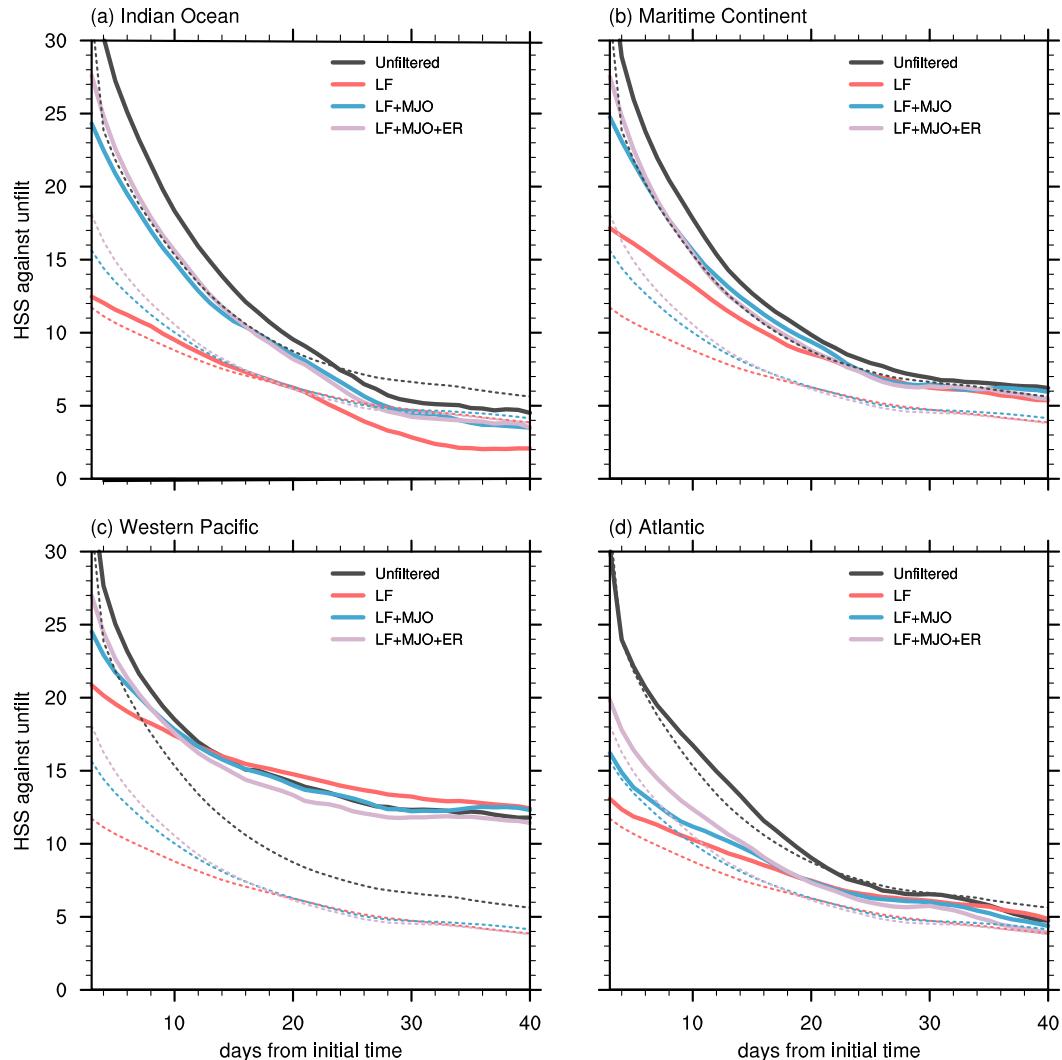


FIG. 7. Heidke skill scores by lead for 7-day running mean forecasts for each of the four regions highlighted by the blue boxes in Figs. 2–6. Dashed lines show the corresponding skill for the entire tropical strip.

Ocean during June–August for phases 1–4. The Maritime Continent (Fig. 8, third row) sees its biggest improvements to skill from the MJO during March–May as well, primarily in phases 3–5. The signals are more mixed in other seasons, with September–November having the largest deductions in skill. The Indian Ocean and the Maritime Continent are also the regions where the relationship between RMM amplitude and MJO skill is clearest.

The MJO generally does not contribute to skill improvements over the western Pacific or the Atlantic (Fig. 8, bottom two rows). After week 1, many phases and seasons shows decreases in skill over both regions. Over the western Pacific, the MJO is most skillful during June–August and least skillful during September–November. These patterns make sense considering

the seasonal cycle of ENSO amplitude (Arguez et al. 2019). ENSO is weakest during boreal summer, so the MJO can play a greater role in rainfall variability. Conversely, ENSO is strongest during boreal fall and winter, which makes it harder for the MJO to contribute to the skill. Over the Atlantic, the MJO produces its largest improvements to skill during March–May. The MJO somewhat improves skill over the Atlantic during September–November, which is a period when the MJO strongly modulates Atlantic tropical cyclone activity (Mo 2000; Klotzbach 2010). Meanwhile, the MJO is the least skillful during December–February.

Figure 9 repeats the analysis for the contribution of ER-filtered anomalies to skill beyond that of LF+MJO. ER waves rarely provide additional skill beyond week 1. ER waves provide the largest additional skill during

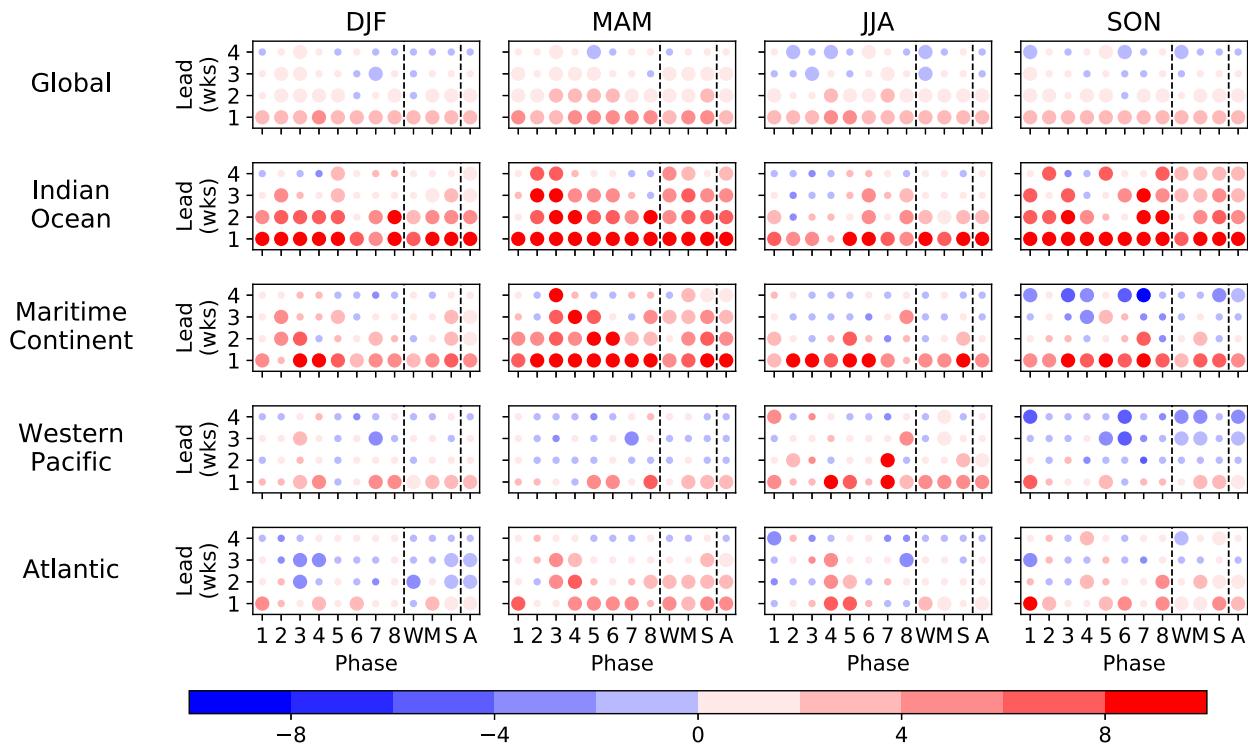


FIG. 8. Difference between HSS for LF+MJO and LF for each region (rows), season (columns), forecast lead, and RMM phase. Red dots indicate that LF+MJO is more skillful, while blue dots show that LF is more skillful. Large dots denote differences that are 95% statistically different from zero. RMM phases are used only when the RMM amplitude is  $>1$ ; W indicates weak MJO ( $RMM < 0.9$ ), M is moderate MJO ( $0.9 \leq RMM < 1.5$ ), S is strong MJO ( $RMM \geq 1.5$ ), and A is for all dates regardless of RMM phase or amplitude.

December–February, particularly over the Indian Ocean, which is a peak time and location for ER wave activity (Roundy and Frank 2004). Perhaps more surprising is that ER waves also contribute skill over the Atlantic during all seasons except for June–August (bottom row). One possible reason for this is that LF and MJO activity are relatively weak over the Atlantic. In addition, ER waves share propagation characteristics with tropical cyclones (Schreck et al. 2011, 2012; Aiyer et al. 2012). These tropical cyclones are themselves modulated by the MJO, so the enhanced skill during September–November may reflect skillful forecasts of tropical cyclone activity (Barnston et al. 2015). However, the regional analyses only include  $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ , which would be south of most tropical cyclones. ER waves also do not provide similar skill in other basins with larger tropical cyclone activity like the Maritime Continent (middle row). ER waves also contribute skill in the Atlantic during December–May. They are generally weaker during that time and region, but they are more equatorially confined.

#### 4. Summary and discussion

This study employs a novel Fourier filtering method (Schreck 2013; Janiga et al. 2018) to identify the

contributions of four subseasonal tropical modes to the CFSv2's skill at predicting rainfall. The initial hypothesis was that the model would be able to predict these envelopes more skillfully than the smaller-scale features that dominate unfiltered rainfall (Tapiador et al. 2019). As in Janiga et al. (2018), this hypothesis was disproven because the filtered skill rarely exceeds that of the unfiltered model forecasts. Even so, the Fourier filtering has two clear uses:

- 1) It is a convenient way to identify sources of skill within the model.
- 2) It can focus the forecaster's attention on the components with the most skill.

In the case of the latter, it allows for straightforward forecast attribution, which enables the forecaster to better communicate both the forecast and associated uncertainty to stakeholders.

The CFSv2 derives most of its skill for week 2 and beyond from simply persisting low-frequency (LF) signals that respond to SST anomalies like ENSO. Comparing the skill of a perfect forecast (Fig. 2a) with the CFSv2 (Fig. 5b) shows that the model harnesses much of the LF potential skill within the oceanic ITCZs. Based on the results of this study, forecasters should place the most

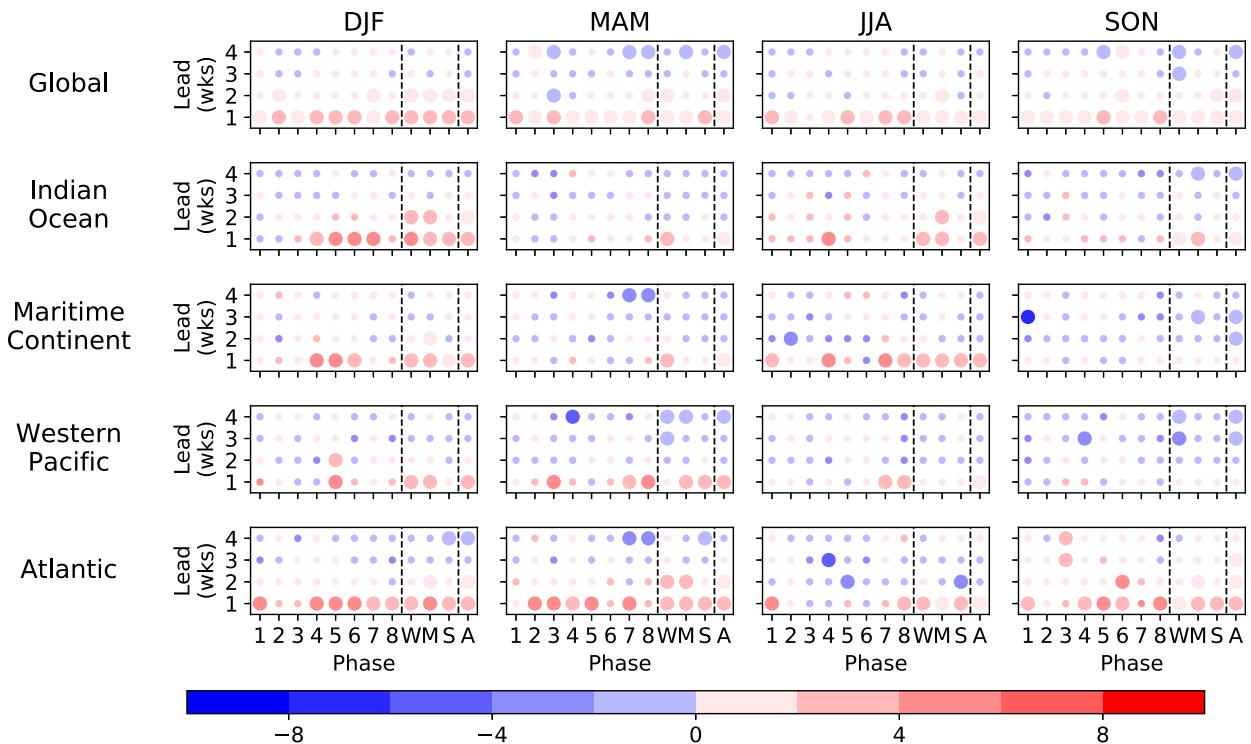


FIG. 9. As in Fig. 8, but for the difference between LF+MJO+ER and LF+MJO.

weight on LF signals from the CFSv2. The MJO can provide additional skill in particular situations. It is generally most skillful over the Indian Ocean and to a lesser degree the Maritime Continent. The MJO signals are generally most reliable during March–May and particularly when the RMM is higher amplitude and in phases 2–5. Meanwhile, the MJO signals should be treated more cautiously during June–August and also over the western Pacific and Atlantic. ER wave signals are generally most skillful during week 1 in December–February and over the Indian Ocean. They are also particularly useful over the Atlantic during September–November.

Most previous studies of model skill for the MJO have focused on the bivariate correlation with the RMM (Gottschalck et al. 2010). For example, Vitart and Molteni (2010) show that the ECMWF can predict the RMM (correlation > 0.5) out to ~20 days. Similarly, Kim et al. (2014) found skill in the VarEPS out to 27 days and the CFSv2 to 21 days. In contrast, this study shows that the week-3 skill from MJO-filtered rainfall is limited to Indian Ocean and to a lesser degree the Maritime Continent. Why are the models so much less skillful in our study? The Fourier filtering in our method smooths both the forecasts and observations in time and longitude, but that seems more likely to improve the apparent skill than detract from it. The more likely explanation is that the RMM is a global index and one that

is heavily biased toward the circulation component of the MJO (Straub 2013). Zonal winds are much smoother in both time and space and thus presumably more predictable than rainfall. Indeed, Janiga et al. (2018) found that models produce substantially better week-3 forecasts for the MJO-filtered zonal winds at 850 and 200 hPa than for OLR. These results should inform future studies that examine model skill for the MJO. Skill scores derived from the RMM alone may hide important model deficiencies in the convection. Comparing Figs. 2 and 5 give some indication of how much room there is for improvement in the rainfall signals associated with the MJO and ER waves in particular. Additionally, future studies should assess the potential of hybrid statistical-dynamical forecast systems that might overcome model shortcomings in forecasting convection.

This study also highlights regions where the MJO and equatorial waves can or cannot contribute to improved subseasonal forecast skill. Using filtered observational data, Fig. 3a suggests that MJO-filtered rainfall anomalies improve the HSS primarily only over the Indian Ocean and the seas around the Maritime Continent. Over the equatorial Pacific, the rainfall anomalies are so strongly modulated by ENSO-related SST variability that the MJO and equatorial waves have little impact there. The CFSv2 largely harnesses the skill associated with these LF signals, but there remains more room for

improvement in its prediction of MJO-related rainfall (cf. Figs. 2d and 5e). The CFSv2's MJO rainfall skill is particularly lacking over the SPCZ and South America.

Kelvin waves remain an interesting challenge for subseasonal dynamics and forecasting. They are synoptic-scale features that are generally identified with narrow filter bands. As a result, they provide only limited potential skill for rainfall. However, individual waves can have impacts and lifespans within the subseasonal scale (Straub et al. 2006). Many of these impacts occur with periods longer than the waves themselves, owing in part to their complex relationship with the MJO. For example, many MJO events can produce subseasonal shifts in convection or circulation that are punctuated by individual Kelvin waves (Schreck and Molinari 2011; Schreck 2015).

The distinctions between Kelvin waves and the MJO also may not be as clear as previously thought. MJO events can transition to more Kelvin-like signatures as they propagate across the Western Hemisphere (Hendon and Liebmann 1994; Straub and Kiladis 2003; Sobel and Kim 2012). The MJO and Kelvin waves also exist on a continuum with some events having hybrid characteristics both in terms of scale and structure (Roundy 2012b,c, 2014). Models with more robust Kelvin waves also tend to have better MJOs (Guo et al. 2015), but model evaluation studies, including the present one, have yet to develop metrics for evaluating the MJO–Kelvin wave continuum. Better evaluation of that continuum may lead to critical improvements in simulating tropical convection.

Finally, this study only evaluated the ensemble mean of the CFSv2. Janiga et al. (2018) performed similar analysis for two other models, but these methods could easily be applied to the full host of subseasonal models to determine the strengths and weaknesses of different models and guide improvements to model physics.

*Acknowledgments.* This research was supported by the NASA Precipitation Measurement Mission (PMM) through Grant NNX16AE33G and by the NOAA OAR Office of Weather and Air Quality (OWAQ) Climate Testbed through Grant NA18OAR4310297. Carl Schreck was also supported by NOAA through the Cooperative Institute for Climate and Satellites–North Carolina under Cooperative Agreement NA14NES432003. Matthew Janiga was supported by the Office of Naval Research and N2N6E through the Navy Earth System Prediction Capability Effort and the Chief of Naval Research through the Naval Research Laboratory Base Program (PE0601153N) and the Office of Naval Research Propagation of Intra-Seasonal Tropical Oscillations (PISTON). CFSv2 reforecasts and calibration climatologies were obtained from NOAA/NCEI.

## REFERENCES

- Aiyyer, A., A. Mekonnen, and C. J. Schreck, 2012: Projection of tropical cyclones on wavenumber–frequency-filtered equatorial waves. *J. Climate*, **25**, 3653–3658, <https://doi.org/10.1175/JCLI-D-11-00451.1>.
- Arguez, A., A. Inamdar, M. A. Palecki, C. J. Schreck, and A. H. Young, 2019: ENSO normals: A new U.S. climate normals product conditioned by ENSO phase and intensity and accounting for secular trends. *J. Appl. Meteor. Climatol.*, **58**, 1381–1397, <https://doi.org/10.1175/JAMC-D-18-0252.1>.
- Barnston, A. G., N. Vigaud, L. N. Long, M. K. Tippett, and J.-K. E. Schemm, 2015: Atlantic tropical cyclone activity in response to the MJO in NOAA's CFS model. *Mon. Wea. Rev.*, **143**, 4905–4927, <https://doi.org/10.1175/MWR-D-15-0127.1>.
- Goswami, B. B., B. Khouider, R. Phani, P. Mukhopadhyay, and A. J. Majda, 2017: Improved tropical modes of variability in the NCEP climate forecast system (version 2) via a stochastic multcloud model. *J. Atmos. Sci.*, **74**, 3339–3366, <https://doi.org/10.1175/JAS-D-17-0113.1>.
- Gottschalck, J., and Coauthors, 2010: A framework for assessing operational Madden–Julian oscillation forecasts: A CLIVAR MJO working group project. *Bull. Amer. Meteor. Soc.*, **91**, 1247–1258, <https://doi.org/10.1175/2010BAMS2816.1>.
- , P. E. Roundy, C. J. Schreck III, A. Vintzileos, and C. Zhang, 2013: Large-scale atmospheric and oceanic conditions during the 2011–12 DYNAMO field campaign. *Mon. Wea. Rev.*, **141**, 4173–4196, <https://doi.org/10.1175/MWR-D-13-00022.1>.
- Guo, Y., D. E. Waliser, and X. Jiang, 2015: A systematic relationship between the representations of convectively coupled equatorial wave activity and the Madden–Julian oscillation in climate model simulations. *J. Climate*, **28**, 1881–1904, <https://doi.org/10.1175/JCLI-D-14-00485.1>.
- Hamill, T. M., and G. N. Kiladis, 2014: Skill of the MJO and Northern Hemisphere blocking in GEFS medium-range reforecasts. *Mon. Wea. Rev.*, **142**, 868–885, <https://doi.org/10.1175/MWR-D-13-00199.1>.
- Hendon, H. H., and B. Liebmann, 1994: Organization of convection within the Madden–Julian oscillation. *J. Geophys. Res.*, **99**, 8073–8083, <https://doi.org/10.1029/94JD00045>.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- Janiga, M. A., C. J. Schreck, J. A. Ridout, M. Flatau, N. P. Barton, E. J. Metzger, and C. A. Reynolds, 2018: Subseasonal forecasts of convectively coupled equatorial waves and the MJO: Activity and predictive skill. *Mon. Wea. Rev.*, **146**, 2337–2360, <https://doi.org/10.1175/MWR-D-17-0261.1>.
- Jie, W., F. Vitart, T. Wu, and X. Liu, 2017: Simulations of the Asian summer monsoon in the sub-seasonal to seasonal prediction project (S2S) database. *Quart. J. Roy. Meteor. Soc.*, **143**, 2282–2295, <https://doi.org/10.1002/qj.3085>.
- Kessler, W. S., 2001: EOF representations of the Madden–Julian oscillation and its connection with ENSO. *J. Climate*, **14**, 3055–3061, [https://doi.org/10.1175/1520-0442\(2001\)014<3055:EROTMJ>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<3055:EROTMJ>2.0.CO;2).
- Kiladis, G. N., K. H. Straub, and P. T. Haertel, 2005: Zonal and vertical structure of the Madden–Julian oscillation. *J. Atmos. Sci.*, **62**, 2790–2809, <https://doi.org/10.1175/JAS3520.1>.
- , M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, <https://doi.org/10.1029/2008RG000266>.

- Kim, H.-M., P. J. Webster, V. E. Toma, and D. Kim, 2014: Predictability and prediction skill of the MJO in two operational forecasting systems. *J. Climate*, **27**, 5364–5378, <https://doi.org/10.1175/JCLI-D-13-00480.1>.
- Klotzbach, P. J., 2010: On the Madden–Julian oscillation–Atlantic hurricane relationship. *J. Climate*, **23**, 282–293, <https://doi.org/10.1175/2009JCLI2978.1>.
- Mo, K. C., 2000: The association between intraseasonal oscillations and tropical storms in the Atlantic basin. *Mon. Wea. Rev.*, **128**, 4097–4107, [https://doi.org/10.1175/1520-0493\(2000\)129<4097:TABIOA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)129<4097:TABIOA>2.0.CO;2).
- Oliver, E. C. J., and K. R. Thompson, 2015: Predictability of the Madden–Julian oscillation index: Seasonality and dependence on MJO phase. *Climate Dyn.*, **46**, 159–176, <https://doi.org/10.1007/s00382-015-2576-3>.
- O’Lenic, E. A., D. A. Unger, M. S. Halpert, and K. S. Pelman, 2008: Developments in operational long-range climate prediction at CPC. *Wea. Forecasting*, **23**, 496–515, <https://doi.org/10.1175/2007WAF2007042.1>.
- Peng, P., A. G. Barnston, and A. Kumar, 2013: A comparison of skill between two versions of the NCEP Climate Forecast System (CFS) and CPC’s operational short-lead seasonal outlooks. *Wea. Forecasting*, **28**, 445–462, <https://doi.org/10.1175/WAF-D-12-00057.1>.
- Roundy, P. E., 2012a: Tracking and prediction of large-scale organized tropical convection by spectrally focused two-step space–time EOF analysis. *Quart. J. Roy. Meteor. Soc.*, **138**, 919–931, <https://doi.org/10.1002/qj.962>.
- , 2012b: The spectrum of convectively coupled Kelvin waves and the Madden–Julian oscillation in regions of low-level easterly and westerly background flow. *J. Atmos. Sci.*, **69**, 2107–2111, <https://doi.org/10.1175/JAS-D-12-060.1>.
- , 2012c: Observed structure of convectively coupled waves as a function of equivalent depth: Kelvin waves and the Madden–Julian oscillation. *J. Atmos. Sci.*, **69**, 2097–2106, <https://doi.org/10.1175/JAS-D-12-03.1>.
- , 2014: Regression analysis of zonally narrow components of the MJO. *J. Atmos. Sci.*, **71**, 4253–4275, <https://doi.org/10.1175/JAS-D-13-0288.1>.
- , and W. M. Frank, 2004: A climatology of waves in the equatorial region. *J. Atmos. Sci.*, **61**, 2105–2132, [https://doi.org/10.1175/1520-0469\(2004\)061<2105:ACOWIT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<2105:ACOWIT>2.0.CO;2).
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, <https://doi.org/10.1175/2010BAMS3001.1>.
- Schreck, C. J., 2013: Comparing MJO diagnostics during DYNAMO. *MJO Field Data and Science Workshop*, Kohala Coast, HI, NCAR/EOL, [https://archive.eol.ucar.edu/projects/dynamo/meetings/2013/mar/abstracts/LargeScale/SchreckDYNAMO2013WorkshopAbstract\\_58032.pdf](https://archive.eol.ucar.edu/projects/dynamo/meetings/2013/mar/abstracts/LargeScale/SchreckDYNAMO2013WorkshopAbstract_58032.pdf).
- , 2015: Kelvin waves and tropical cyclogenesis: A global survey. *Mon. Wea. Rev.*, **143**, 3996–4011, <https://doi.org/10.1175/MWR-D-15-0111.1>.
- , 2016: Convectively coupled Kelvin waves and tropical cyclogenesis in a semi-Lagrangian framework. *Mon. Wea. Rev.*, **144**, 4131–4139, <https://doi.org/10.1175/MWR-D-16-0237.1>.
- , and J. Molinari, 2011: Tropical cyclogenesis associated with Kelvin waves and the Madden–Julian oscillation. *Mon. Wea. Rev.*, **139**, 2723–2734, <https://doi.org/10.1175/MWR-D-10-05060.1>.
- , —, and K. I. Mohr, 2011: Attributing tropical cyclogenesis to equatorial waves in the western North Pacific. *J. Atmos. Sci.*, **68**, 195–209, <https://doi.org/10.1175/2010JAS3396.1>.
- , —, and A. Ayyer, 2012: A global view of equatorial waves and tropical cyclogenesis. *Mon. Wea. Rev.*, **140**, 774–788, <https://doi.org/10.1175/MWR-D-11-00110.1>.
- Sobel, A. H., and D. Kim, 2012: The MJO–Kelvin wave transition. *Geophys. Res. Lett.*, **39**, L20808, <https://doi.org/10.1029/2012GL053380>.
- Sooraj, K. P., H. Annamalai, A. Kumar, and H. Wang, 2012: A comprehensive assessment of CFS seasonal forecasts over the tropics. *Wea. Forecasting*, **27**, 3–27, <https://doi.org/10.1175/WAF-D-11-00014.1>.
- Straub, K. H., 2013: MJO initiation in the real-time multivariate MJO index. *J. Climate*, **26**, 1130–1151, <https://doi.org/10.1175/JCLI-D-12-00074.1>.
- , and G. N. Kiladis, 2003: Interactions between the boreal summer intraseasonal oscillation and higher-frequency tropical wave activity. *Mon. Wea. Rev.*, **131**, 945–960, [https://doi.org/10.1175/1520-0493\(2003\)131<0945:IBTBSI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0945:IBTBSI>2.0.CO;2).
- , —, and P. E. Ciesielski, 2006: The role of equatorial waves in the onset of the South China Sea summer monsoon and the demise of El Niño during 1998. *Dyn. Atmos. Oceans*, **42**, 216–238, <https://doi.org/10.1016/j.dynatmoce.2006.02.005>.
- Tapiador, F. J., R. Roca, A. Del Genio, B. Dewitte, W. Petersen, and F. Zhang, 2019: Is precipitation a good metric for model performance? *Bull. Amer. Meteor. Soc.*, **100**, 223–233, <https://doi.org/10.1175/BAMS-D-17-0218.1>.
- Ventrice, M. J., and C. D. Thorncroft, 2013: The role of convectively coupled atmospheric Kelvin waves on African easterly wave activity. *Mon. Wea. Rev.*, **141**, 1910–1924, <https://doi.org/10.1175/MWR-D-12-00147.1>.
- , —, and M. A. Janiga, 2012a: Atlantic tropical cyclogenesis: A three-way interaction between an African easterly wave, diurnally varying convection, and a convectively coupled atmospheric Kelvin wave. *Mon. Wea. Rev.*, **140**, 1108–1124, <https://doi.org/10.1175/MWR-D-11-00122.1>.
- , —, and C. J. Schreck, 2012b: Impacts of convectively coupled Kelvin waves on environmental conditions for Atlantic tropical cyclogenesis. *Mon. Wea. Rev.*, **140**, 2198–2214, <https://doi.org/10.1175/MWR-D-11-00305.1>.
- Vitart, F., and F. Molteni, 2010: Simulation of the Madden–Julian oscillation and its teleconnections in the ECMWF forecast system. *Quart. J. Roy. Meteor. Soc.*, **136**, 842–855, <https://doi.org/10.1002/qj.623>.
- Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, **56**, 374–399, [https://doi.org/10.1175/1520-0469\(1999\)056<0374:CCEWAO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2).
- , and K. M. Weickmann, 2001: Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Mon. Wea. Rev.*, **129**, 2677–2694, [https://doi.org/10.1175/1520-0493\(2001\)129<2677:RTMAPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2677:RTMAPO>2.0.CO;2).
- , and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, [https://doi.org/10.1175/1520-0493\(2004\)132<1917:AARMMI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2).
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 648 pp.
- Yoneyama, K., C. Zhang, and C. N. Long, 2013: Tracking pulses of the Madden–Julian oscillation. *Bull. Amer. Meteor. Soc.*, **94**, 1871–1891, <https://doi.org/10.1175/BAMS-D-12-00157.1>.
- Zhang, C., 2013: Madden–Julian oscillation: Bridging weather and climate. *Bull. Amer. Meteor. Soc.*, **94**, 1849–1870, <https://doi.org/10.1175/BAMS-D-12-00026.1>.